Authors' response to reviewer 3

We are thankful for having the manuscript "The Linear Link: Deriving Age-Specific Death Rates from Life Expectancy" reviewed. We appreciate the time and effort dedicated to providing feedback on our manuscript and are grateful for the insightful comments and valuable improvements to our paper. Please see below, in blue, for a point-by-point response to all comments and concerns.

Open Review

( ) I would not like to sign my review report
(x) I would like to sign my review report

English language and style

( ) Extensive editing of English language and style required
( ) Moderate English changes required
( ) English language and style are fine/minor spell check required
(x) I don't feel qualified to judge about the English language and style

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	( )	( )	(x)	( )
Is the research design appropriate?	( )	( )	( )	(x)
Are the methods adequately described?	( )	( )	(x)	( )
Are the results clearly presented?	( )	( )	(x)	( )
Are the conclusions supported by the results?	( )	( )	(x)	( )

Comments and Suggestions for Authors

My review is based on a general comment and further particular comments.

As a general comment, this paper proposes a new relational model. This model is inspired by indirect estimation techniques applied in demography, which can be used to estimate full life tables at any point in time, based on a given value of life expectancy at birth or life expectancy at other age. This new model is useful for age patterns of mortality in poor-data countries; nevertheless, authors fit the model using the death rates computed using death counts and population exposed to the risk of death in the calendar year for the female populations of the England &Wales, France, Sweden and USA available in the HMD.

In my opinion, the main drawbacks are that the new model should test linearity in the relation between age-specific death mortality and life expectancy, and prediction should be using confidence intervals.

Authors` response: Thank you for pointing at these drawbacks. We have now emphasized in Figures 1 and 2 the existent strong linear relations between life expectancies and death rates, Pearson correlation coefficients are now included there. Similarly, the prediction intervals included in Figure 7 are briefly described in the methods section.

The authors should also state three crucial points in the Abstract and in the Introduction of the paper:

1. How the existence of their findings in mortality data could help practitioners to improve the understanding of mortality data features.
2. The main contributions of the paper to obtain life tables comparing to previous papers. Authors should describe poor-data where life expectancy at birth is known but not death rates.
3. If the life tables obtained are accurate enough to evaluate pension systems sustainability.

We appreciate these suggestions and have included elements of 1, 2 and 3 above in the abstract and introduction.

It would be interesting to know how to implement the model in MortalityLaws R-package; now, it is not in vignettes, maybe authors can add an Appendix.

Authors` response: This is an interesting suggestion. At this point there are no plans to extend the MortalityLaws R-package. But we will consider that possibility in the near future. As you correctly noticed we have created the open source MortalityEstimate R-packge available on Github to host such models.

On the other hand, the **Reproducible research** section is better to include URL https://github.com/mpascariu/MortalityEstimate instead of the general URL https://github.com.

Authors` response: In the final version of the article the correct link will be available.

Finally, the author(s) should discuss their approach's real advantage compared to the existing ones. Results should be revised with more detail, and some conclusions or practical interpretations about them should be indicated in the Conclusions.

For example, Do the results depend on the country?

Authors` response: We thank the reviewer for this suggestion, which echoes some of the comments of the other reviewers. Action: We have now highlighted with more detail the advantages of the linear link model respect to others: in particular, we discuss the advantage of using the life expectancy in our equation (1) which is justified by the strong correlations found in our Figures (1) and (2) between life expectancy and death rates; as opposed to Ševčíková et al. (2016) that use life expectancy as a constrain to estimate parameters, we include life expectancy as part of the initial estimation creating, we consider, a straightforward procedure; additionally we

presented and discuss the results of our model in various countries and tested the accuracy of the model in relation to the Lee-Carter model.

Indeed, we presented results for 4 populations where we noticed a different degree of model performance given by the mortality specificities of those populations. We hope that our openness of sharing the code of the programs used and the free availability of the HMD will invite readers to try to replicate our results for the countries presented and for any other population. We would be delighted to hear back from users on advantages and disadvantages that they find along the way and we will leave an open channel to correspond with them.

Are there different results for different age ranges?

Authors` response: Using the Linear-Link model one can reconstruct the mortality experience over the entire age range at a given point in time. It goes without saying that the mortality estimates for each particular age might differ from the true experience with a higher of lower degree of error and our final proposed forecasts include prediction intervals which differ from age to age capturing some of this variability. Action: We have included this in the discussion.

How does this result help actuaries in pricing?

Authors` response: The life insurance instruments, pension valuations, longevity transfer deals and other technical aspects in the field of actuarial depend heavily on accurate demographic data. The present method offers a solution of obtaining such reliable estimates when life expectancy at birth is targeted in future projections. Action: we have highlighted this in the discussion section.

Can models be used in a portfolio?

Authors` response: In the manuscript we have shown how the method is applied to national population, however in the same way it can be used to modelized subpopulation given by insurance portfolios. This would be a perfect example of a population that is subject to an elevated level of life expectancy compared with the general population and where data is scarce and short in duration. The short answer is yes! Action: in the discussion we mention this possible use of the linear-link model.

Particular comments:

Pag 2, lines 57-65. The author(s) should discuss the real advantage of their approach compared to each relational model. Where is the research gap?

Authors` response: The novelty of our approach is the reliance on the single most comprehensive demographic measure to measure quality of life in a population i.e. life expectancy. This idea combined with a proven Lee-Carter like method for deriving age patterns of mortality demonstrated in the presented results would make a great scientific addition to the field. Action: we wrote "Our model recovers the entire age profile of mortality in a population, based on the strong correlations between a single longevity measure, namely life expectancy, and age-specific

death rates. Combining these correlations in the Lee-Carter (1992) methodology makes the proposed algorithm appealing to be used in forecasting practice."

Pag 2, lines 74-75. The suggestion is to show poor data countries with life known expectancy at birth that need to predict death rates. Therefore, an algorithm that derives a life table based only on life expectancy at birth can also be widely used in forecasting practice.

Authors` response: We have deferred this comment for the discussion where we include how this can be applied in data poor settings. Action: we wrote "For data-poor populations with a desired level of life expectancy the age-profile of mortality can be attained by borrowing the model's parameters from a country with similar profile."

Pag 2, lines 16-18. Which are the main conclusions of these papers on relational models? Which countries are studied? In my opinion, the authors enumerate different relational models, which is not enough to justify another paper.

Authors` response: The presented articles on relational models are indicated in order to acknowledge the existing contributions to the field mentioning their main strategy of estimating mortality. Comparing many relational models is out of the scope of this manuscript, and our goal is mainly to contrast the new proposal with the closest type of model, namely the Lee-Carter and the Ševčíková et al. (2016) model. Action: As mentioned above, we have now highlighted with more detail the advantages of the linear link model respect to others: in particular, we discuss the advantage of using the life expectancy in our equation (1) which is justified by the strong correlations found in our Figures (1) and (2) between life expectancy and death rates; as opposed to Ševčíková et al. (2016) that use life expectancy as a constrain to estimate parameters, we include life expectancy as part of the initial estimation, creating we consider, a straightforward procedure; additionally we presented and discuss the results of our model in various countries and tested the accuracy of the model in relation to the Lee-Carter model.

Pag 2, line 73. Please, quote the web and reference of the HMD database.

Authors` response: In the revised manuscript we have improved the citation in the refence table including the URL. The database is cited in line 67.

Pag 2, lines 88-89. Authors should justify the countries' selection based on actuaries or demographers' interests, countries with different demographic and economic evolution, and especially the data-poor context. Results can change with these choices.

Authors` response: We agree that the accuracy of the results can change form country to country. This being a methodological paper we choose to present show populations where the method can be tested, and its accuracy measured. Action: we wrote " the population selection was based on different degree of model performance given by the mortality specificities of those populations, for example old age mortality in the HMD is often subject to diverse correction procedures and modelling depending on the country (Wilmoth et al. 2007).

Pag 2, line 91 only mortality data between age 0 and 100 are used in the model, but Figure 1 shows age 110. In this figure, sparse data corresponds to age 10 or 20, not only to ages higher to 100. It is worth mentioning that just a graph is subjective to justify linearity; authors should use the Pearson correlation coefficient.

Authors` response: We are thankful for this comment and we acknowledge that the linearity across some ages can be challenged. This is the case for ages above 100 and for ages between 10 and 20 where the transition from a decreasing mortality to an Gompertzian mortality takes place. Action: We have computed the Pearson correlation coefficient for each age group in Figure 1 and Figure 2 in order to give a better understability of the linearity between life expectancy and death rates.

Pag 3, line 101 εxt is not a parameter. Authors

Authors` response: Typo corrected.

Pag 3, line 112 Authors should not talk about strong linear trends based on Figures 1 and 2 that depends on graph scale.

Authors` response: As suggested by the reviewer we have computed the Pearson correlation coefficient for each age group in Figure 1 and Figure 2 to emphasize the linearity between life expectancy and death rates and also make a distinction between the linearity at adult ages and the one exhibited at young or very old ages.

Pag 4. May the authors provide parameter significance?

Authors` response: The meaningful interpretation of the parameters functions is in the same direction as in the Lee-Carter model. Action: we wrote "Indeed, if one sets the parameters $b_x log e_{\theta,t} = a_x, v_x = b_x, and\ k = k_t$, we obtain the LC model. Interpretation of the parameters are then similar with $b_x log e_{\theta,t}$ a standard age profile, v_x the age-specific improvements in mortality, and k the amount of average mortality improvement."

Pag 4 To gain precision in the fitting authors should use confidence intervals for age-specific death rates.

Authors` response: Confidence intervals are presented in Figure 7. We are making our source code open to give access to readers to detailed numerical results that cannot be included in the manuscript without using large tables.

Pag 4. Figure 2 label of y-axis is not life expectancy at birth; it is life expectancy at age 65. Why have not authors done the model excluding young ages? Maybe it more interesting for actuaries.

Authors` response: The model can be adapted for ages 65+ by simplifying it due to a more regular mortality pattern. We choose to estimate it for the entire age range in order to test the most challenging scenario.

Pag 5, line 142, which Appendix A or B?

Authors` response: Appendix A. Now clearly specified in the manuscript.

Pag 5, line 151 see Appendix B. On the other hand, minor discrepancies in some examples do not justify the use of least squares over maximum likelihood.

Authors` response: We fully agree with the reviewer. For this reason, we have provided both options, the least squares and the maximum likelihood in our R code for the user to be able to take an educated decision in fitting the model.

Pag 7, Please, the expression of relative error should be given; results seem very high, especially in Figure 6.

Authors` response: The formula is (O - E)/E * 100, where the O and E denotes the observed and estimated log death rates. The accuracy measure at very young and very old ages display high values because the observed data is very volatile in these age ranges. Since the output given by the Linear-Link model is in general smooth the high errors are to be expected. Action: We included details about the formula in the caption of Figure 6.

Pag 14-16. Please, revise and update references.

Authors` response: Done.

Typing errors

Pag 1: There is a mistake in the reference? which means a lack of this reference.

Authors response: Thank you for noticing this error. It has caused by the manuscript compilation procedure. In the improved paper this has been corrected.